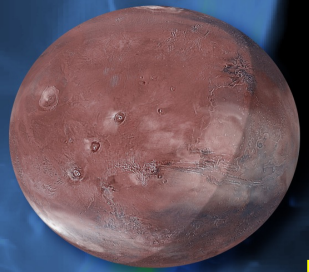
National Aeronautics and Space Administration



NASA Center for Climate Simulation – Change is Constant

Laura Carriere, NCCS HPC Lead

HPC User Forum – Spring 2022

NCCS
NASA CENTER FOR CLIMATE SIMULATION
HIGH-PERFORMANCE SCIENCE

- **NASA Center for Climate Simulation**
  - One of two NASA Supercomputer facilities, located at Goddard Space Flight Center in Greenbelt, MD
  - Focus on climate research
- Collaboration with the NASA Advanced Supercomputing Division at NASA Ames

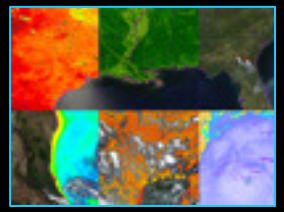
NASA Center for Climate Simulation Facility

About the NCCS

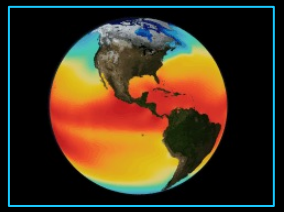
Part of Computational and Information Sciences and Technology Office (CISTO)

High End Networks and Information Technology Security Group

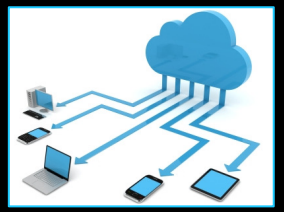
NCCS – High Performance Computing (HPC)



Data Science Group, Large Scale Analytics, Artificial Intelligence, and Machine Learning



Scientific Visualization Studio (SVS) – *funded outside Scientific Computing*



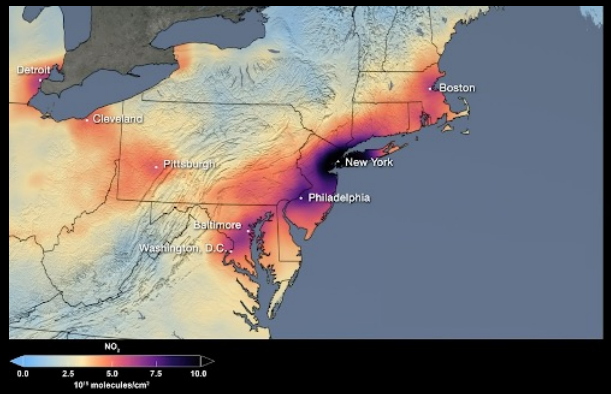
Science Managed Cloud Environment (SMCE)

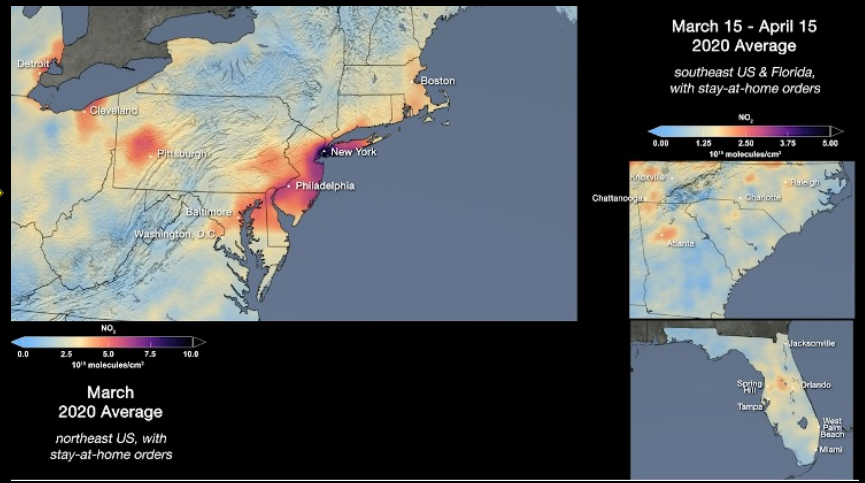
User Community is the Goddard Sciences and Exploration Directorate, e.g.:

- Global Modeling and Assimilation Office (GMAO)
- Goddard Institute for Space Studies (GISS)
- Hydrological Sciences Laboratory
- Lots of others

GEOS-CF COVID-19 NO_2 Emissions

GEOS Composition Forecast Model was run at 25-kilometer resolution on 3,416 cores of the NCCS Discover supercomputer and produces approximately 300 gigabytes of output per day found reductions of 20%-50%





Credit: Trent Schindler (SVS)

March 2015-2019 Average

March 2020

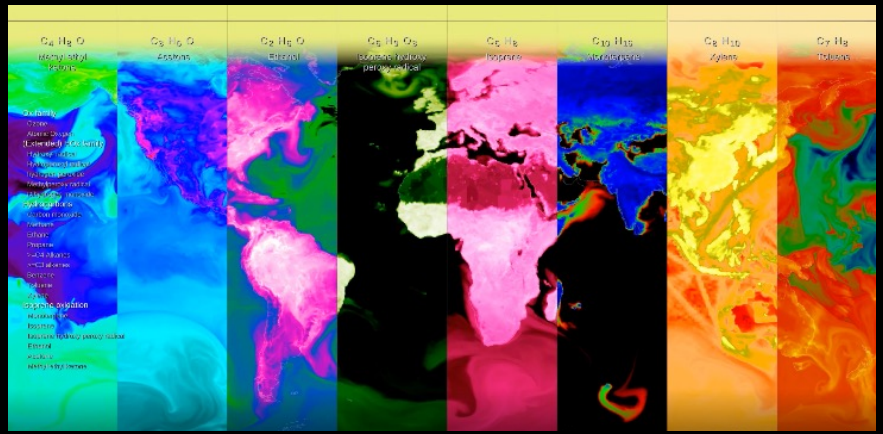
GEOS Atmospheric Chemistry

GEOS-CF modeled 240 chemical species over 3 months

Incorporates the latest scientific understanding of the physics and chemistry that guide the formation of ozone, along with measurements from satellites and other instrument platforms.

These chemicals undergo rapid changes as they are being emitted by natural and anthropogenic activities, transported by prevailing winds and vertical lifting motions, deposited to the surface, or chemically transformed.

Visualization shows 96 species, e.g. Ozone, Hydrocarbons, Aerosols, and Halogens

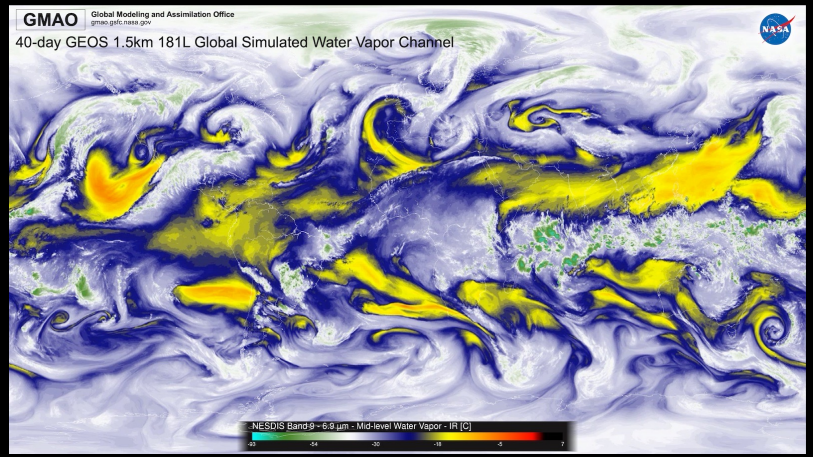


Credit: Greg Shirah, SVS

GMAO GEOS DYAMOND

DYAMOND Phase II – An international project to study northern hemisphere winter simulations with atmosphere-ocean coupling.

The GMAO improved the representation of convection, precipitation, aerosols, carbon, and chemistry in the NASA GEOS model. Runs required just under 40,000 Intel Skylake cores and produced 3.3 PB of output

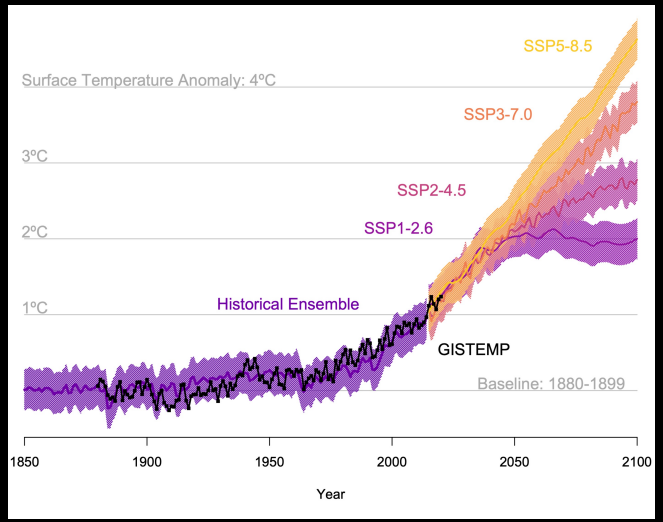


Credit: William Putman, NASA GSFC

A snapshot of mid-level simulated water vapor from the GEOS 1.5-km global DYAMOND Phase II 40-day simulation. This image details the presence of dry (yellow-orange) and wet (white-green) air in the middle levels of the troposphere.

GISS CMIP6

GISS's CMIP6 runs have consumed approximately 200 million core hours and the NCCS hosts over 150 TB of GISS CMIP6 data, available through NCCS and ESGF.



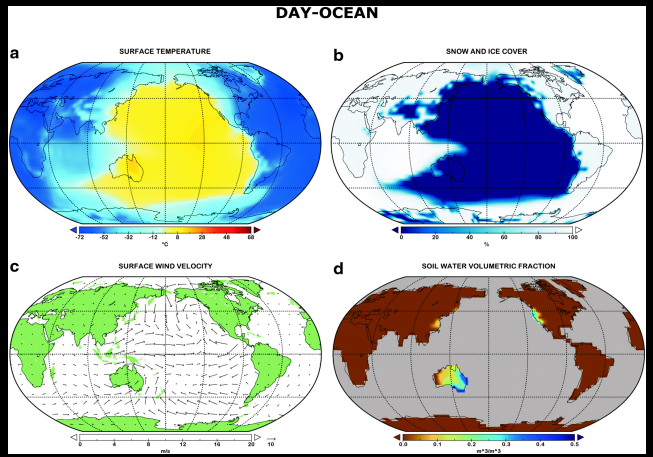
Credit: NASA GISS

GISS researchers ran several hundred paleoclimate (distant past), historical (starting in 1850), carbon cycle, and future simulations. In the GISS CMIP6 simulations, historical drivers of emissions and greenhouse gas concentrations produce responses (purple line) that match well to GISTEMP surface temperature observations (black line) going back to 1880.

GISS in now using their ModelE 3 — with 111 km grid resolution and 106 levels which uses a cubed-sphere atmospheric dynamical core and grid as well as new "moist physics".

GISS Exoplanets

Ran 43 ROCKE-3D simulations using 44 cores and producing 3.3 TB of output.
Found that under a wide range of conditions, Proxima Centauri b can host liquid water on much of its surface, potentially raising its prospects for harboring living organisms
Simulations assumed an atmosphere and an ocean, 11-day year, tidally locked

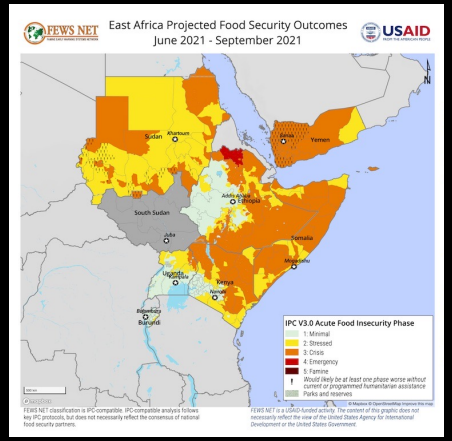


Atmosphere simulation of Proxima Centauri b, using Earth's continental configuration.

Credit: Del Genio et al.D

Excellent example of NASA's ability to do interdisciplinary science – results may guide researchers to the most promising exoplanet candidates for James Webb Space Telescope observations

Other NCCS Science Results



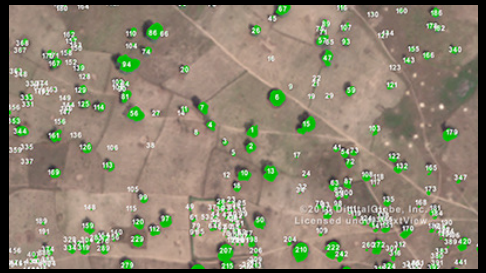
Credit: Abheera Hazra, Kimberly Slinski, NASA

Hydrological Forecast System

NHyFAS is uses to produce monthly drought and flood potential monitoring and seasonal forecast products in a partnership with USAID and DoD

Satellite Tree Enumeration

HPC is used to generate DEMs from very high-resolution satellite imagery.
AI/ML is used to identify and count trees and shrubs

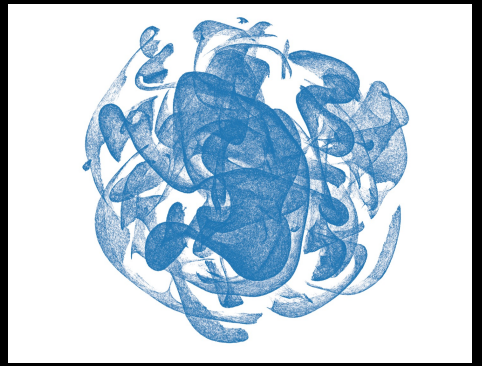


Credit: Greg Shirah, Compton Tucker, NASA

Change is ability to use HPC and AI/ML

ML and TESS

HPC was used to generate 70M light curves from TESS
AI/ML was used to identify binary systems and a sextuply star system



Credit: Brian Powell

Discover

- Intel based 125,000 cores, 12 NVIDIA A100 nodes
- IB and OPA Fabrics
- 63 PB online storage, 110 PB on tape
- Slurm scheduler
- Expansion is done in scalable units – working on SCU17 - AMD Milan

SCU14

Discover Base Unit (2006)

- 130 nodes, 4 cores/node
- 520 total cores
- 3.3 TF
- $515K/TF

Intel Dempsey

Discover SCU17 (2022) – 1 chassis

- 4 nodes, 128 cores/node
- 512 total cores
- Chassis – 16.4 TF
- Node – 4.1 TF
- $4K/TF

AMD Milan

Mass Storage Tape Library

110 PB but difficult to access for today's AI/ML workflows

CSS

Centralized NASA data – 45 PB

Shared with all NCCS environments

Data Management Plans

Prism

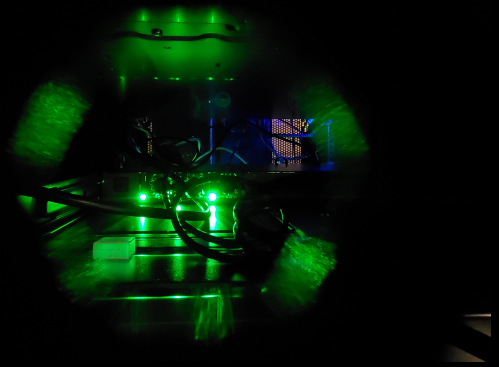
GPU cluster for AI/ML

22 systems 4x NVIDIA V100 GPUs

1 DGX with 4x NVIDIA A100 GPUs

ADAPT/Explore

On-premise Cloud environment



Started as proof of concept with repurposed Discover nodes, is now built from new AMD nodes and repurposed Intel nodes

Explore – OpenStack On-Premise Cloud

Rationale – Complimentary environment to HPC (simulations, MPI, shared parallel filesystem, managed software environment), Explore provides access to large data sets mostly analysis codes, ability to build customized software environments for a group or user

Architecture – Multiple availability zones, with redundant storage and network access, across two different datacenters at Goddard, PIV integration of Keystone/Horizon

Capacity – Total cores 8768, 72 TB RAM, 309 TF

Features – IaaS, PaaS

Implements the OpenStack Python API to provide easier access for scripting and automating cloud activities

Shared storage with HPC, Data Services environments through CSS allows users to compute in both environments on the same data

SMCE/AWS

Rationale – A managed cloud provides the framework for NASA users to quickly get started

Objectives - Facilitate and accelerate collaboration, rapid prototyping, and science research using commercial cloud computing resources

Benefits – Provides security layer for FISMA Low, financial management, cloud training, ease of collaboration with external scientists

Solutions – Gitlab using Kubernetes for CI/CD runners, Jupyter ecosystem, AI/ML workflows, HPC parallel clusters

Data – NASA is moving lots of data to AWS

HPC Research stays on-premise due to cost of experimentation

Future System Challenges

Still power constrained – Discover drops to 2/3 capacity during SCU integrations, need to identify and migrate code to the cloud during integrations

On-Premise cloud and HPC coexistence, including orchestration, containers, graphical interfaces, OS choice

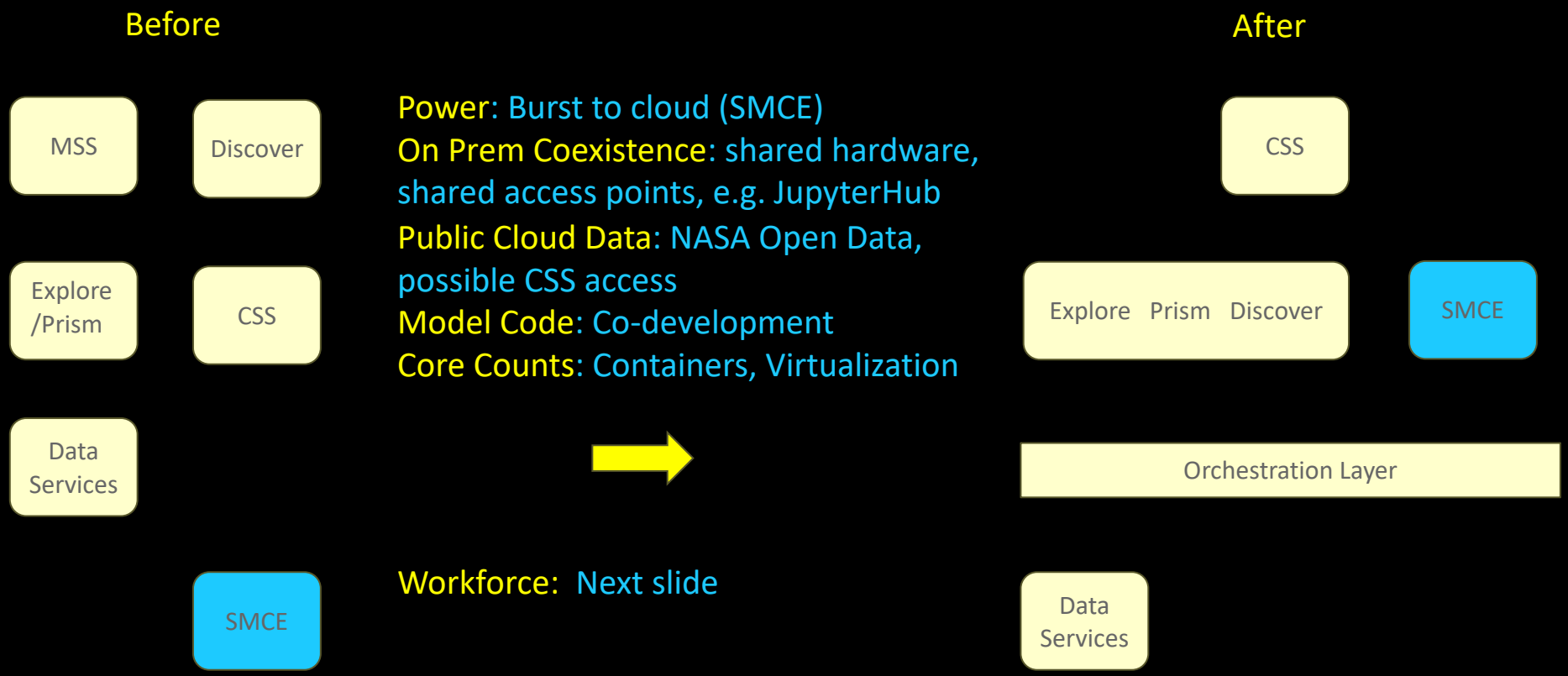
Public Cloud coexistence, including data location

Climate Model Code

- Doesn't easily scale – legacy memory management decisions
- Doesn't port easily to GPUs or integrate AI – progress here
- Efficiency questions as chips become denser (AMD's 96 and 128 core chips)

Staffing – combination of HPC and Cloud expertise

Future System



Workforce

Campus Research Computing Consortium (CaRCC) Career Arcs Working Group

Based on the Research Computing and Data (RCD) Capabilities Model to describe five "facings" – Research, Data, Software, Systems, and Strategy and Policy

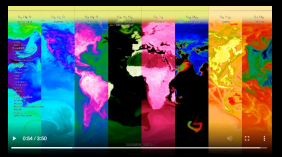
Two surveys – RCD Professionalization (Job titles, roles, salaries) and RCD Career Arc (recruitment, advancement, retention)

Link to paper when published: https://carcc.org/career-arcs/

Quick take away:

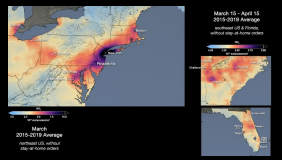
- Women were more likely than men to believe referrals matter
- Men were more likely than women to rely on referrals for hiring and promotion
- Referrals matter – one way each of us can act is to make more referrals of deserving women and people of color for jobs into which they can grow

Links to Visualizations



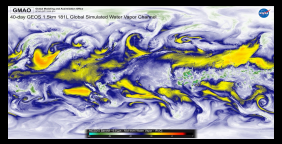
GEOS Composition Forecast (GEOS-CF)

https://svs.gsfc.nasa.gov/4754



GEOS-CF COVID-19 Emissions

https://svs.gsfc.nasa.gov/31142 https://svs.gsfc.nasa.gov/13753



DYAMOND Run

https://www.nas.nasa.gov/SC21/research/project32.html

NCCS Science Stories:

https://www.nccs.nasa.gov/news-events/nccs-highlights/